



WORKSHEET

Data Representation Question – Cumulative *Frequency Graphs*

Question Description: Every year in your Statistics 1 exam there is a question on Representing Data. Bellow is all the exam questions when they ask you to represent data by drawing a “cumulative frequency graph”.

QUESTIONS

Q1 s2015_p63

Seventy samples of fertiliser were collected and the nitrogen content was measured for each sample. The cumulative frequency distribution is shown in the table below.

Nitrogen content	≤ 3.5	≤ 3.8	≤ 4.0	≤ 4.2	≤ 4.5	≤ 4.8
Cumulative frequency	0	6	18	41	62	70

- (i) On graph paper draw a cumulative frequency graph to represent the data. [3]
- (ii) Estimate the percentage of samples with a nitrogen content greater than 4.4. [2]
- (iii) Estimate the median. [1]

Q2 s2016_p61

The amounts spent by 160 shoppers at a supermarket are summarised in the following table.

Amount spent (\$ x)	$0 < x \leq 30$	$30 < x \leq 50$	$50 < x \leq 70$	$70 < x \leq 90$	$90 < x \leq 140$
Number of shoppers	16	40	48	26	30

- (i) Draw a cumulative frequency graph of this distribution. [4]
- (ii) Estimate the median and the interquartile range of the amount spent. [3]
- (iii) Estimate the number of shoppers who spent more than \$115. [2]
- (iv) Calculate an estimate of the mean amount spent. [2]

Q3 w2016_p63

The tables summarise the heights, h cm, of 60 girls and 60 boys.

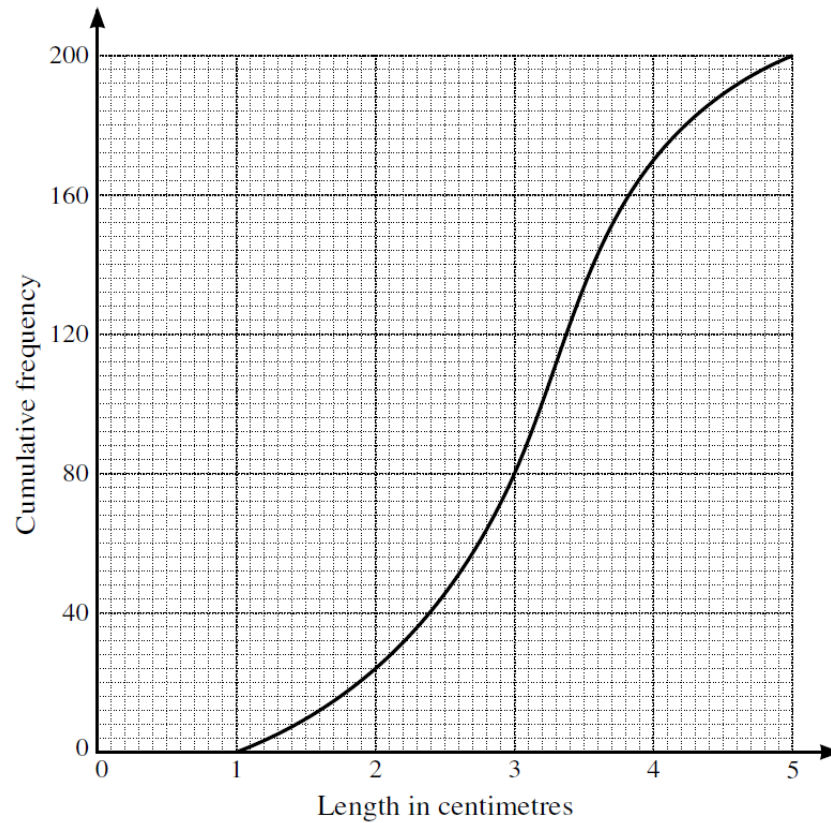
Height of girls (cm)	$140 < h \leq 150$	$150 < h \leq 160$	$160 < h \leq 170$	$170 < h \leq 180$	$180 < h \leq 190$
Frequency	12	21	17	10	0

Height of boys (cm)	$140 < h \leq 150$	$150 < h \leq 160$	$160 < h \leq 170$	$170 < h \leq 180$	$180 < h \leq 190$
Frequency	0	20	23	12	5

- (i) On graph paper, using the same set of axes, draw two cumulative frequency graphs to illustrate the data. [4]
- (ii) On a school trip the students have to enter a cave which is 165 cm high. Use your graph to estimate the percentage of the girls who will be unable to stand upright. [3]
- (iii) The students are asked to compare the heights of the girls and the boys. State one advantage of using a pair of box-and-whisker plots instead of the cumulative frequency graphs to do this. [1]

Q4 s2017_p62

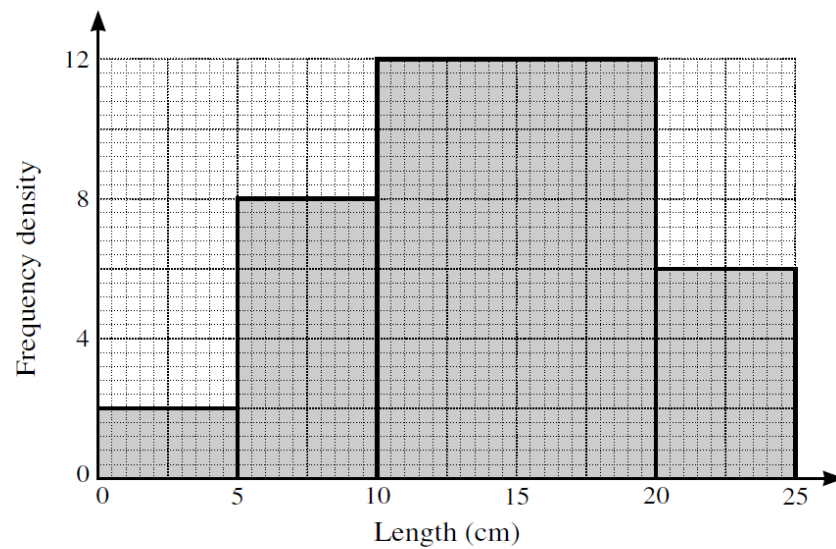
Anabel measured the lengths, in centimetres, of 200 caterpillars. Her results are illustrated in the cumulative frequency graph below.



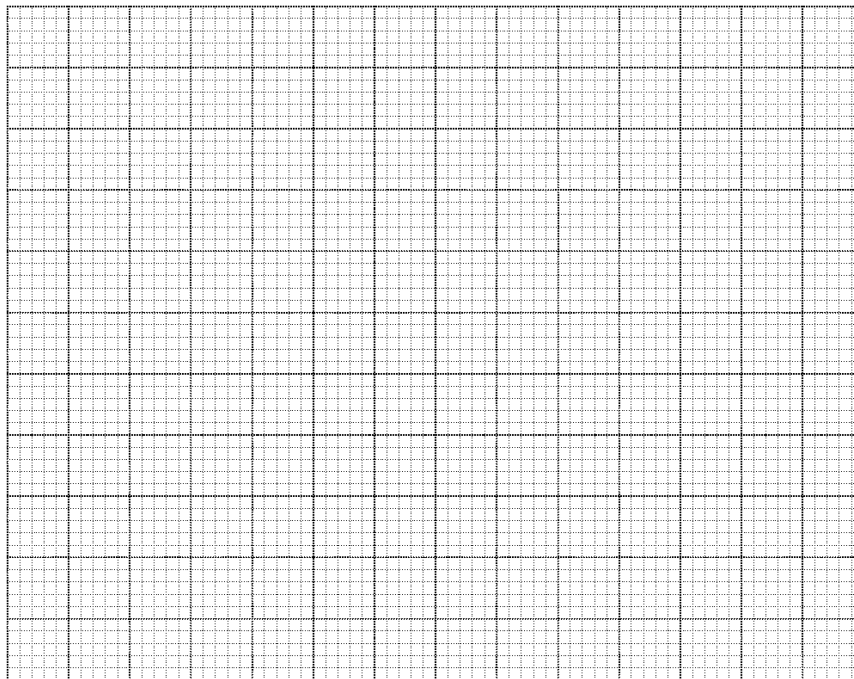
- (i) Estimate the median and the interquartile range of the lengths. [3]
- (ii) Estimate how many caterpillars had a length of between 2 and 3.5 cm. [1]
- (iii) 6% of caterpillars were of length l centimetres or more. Estimate l . [2]

Q5 s2017_p63

The following histogram represents the lengths of worms in a garden.



- (i) Calculate the frequencies represented by each of the four histogram columns. [2]
- (ii) On the grid on the next page, draw a cumulative frequency graph to represent the lengths of worms in the garden. [4]



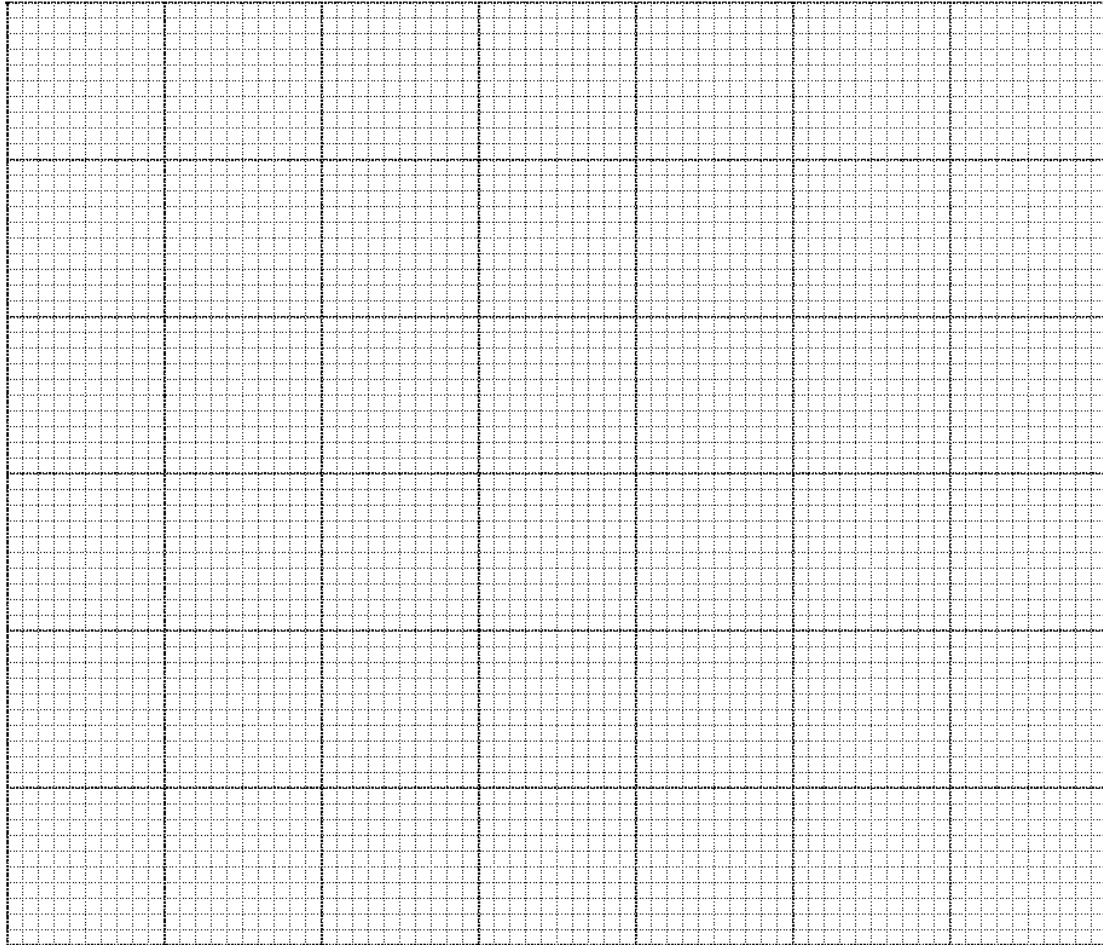
- (iii) Use your graph to estimate the median and interquartile range of the lengths of worms in the garden. [3]
- (iv) Calculate an estimate of the mean length of worms in the garden. [2]

Q6 w2017_p61

The time taken by a car to accelerate from 0 to 30 metres per second was measured correct to the nearest second. The results from 48 cars are summarised in the following table.

Time (seconds)	3 – 5	6 – 8	9 – 11	12 – 16	17 – 25
Frequency	10	15	17	4	2

- (i) On the grid, draw a cumulative frequency graph to represent this information. [3]



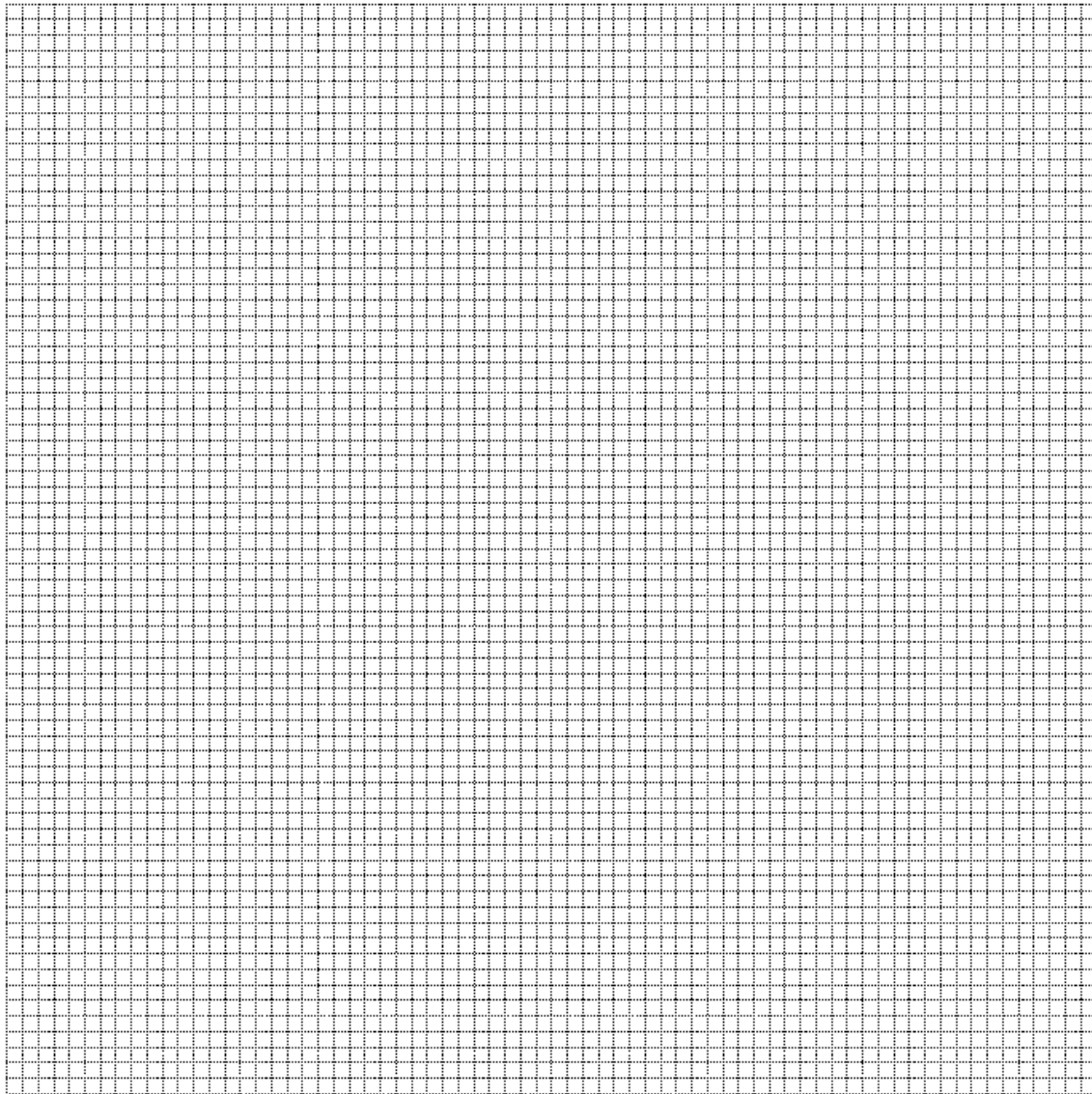
- (ii) 35 of these cars accelerated from 0 to 30 metres per second in a time more than t seconds. Estimate the value of t . [2]

Q7 w2017_p62

The circumferences, c cm, of some trees in a wood were measured. The results are summarised in the table.

Circumference (c cm)	$40 < c \leq 50$	$50 < c \leq 80$	$80 < c \leq 100$	$100 < c \leq 120$
Frequency	14	48	70	8

- (i) On the grid, draw a cumulative frequency graph to represent the information. [3]



- (ii) Estimate the percentage of trees which have a circumference larger than 75 cm. [2]

Q8 w2018_p61

The daily rainfall, x mm, in a certain village is recorded on 250 consecutive days. The results are summarised in the following cumulative frequency table.

Rainfall, x mm	$x \leq 20$	$x \leq 30$	$x \leq 40$	$x \leq 50$	$x \leq 70$	$x \leq 100$
Cumulative frequency	52	94	142	172	222	250

(i) On the grid, draw a cumulative frequency graph to illustrate the data.

[2]



(ii) On 100 of the days, the rainfall was k mm or more. Use your graph to estimate the value of k .

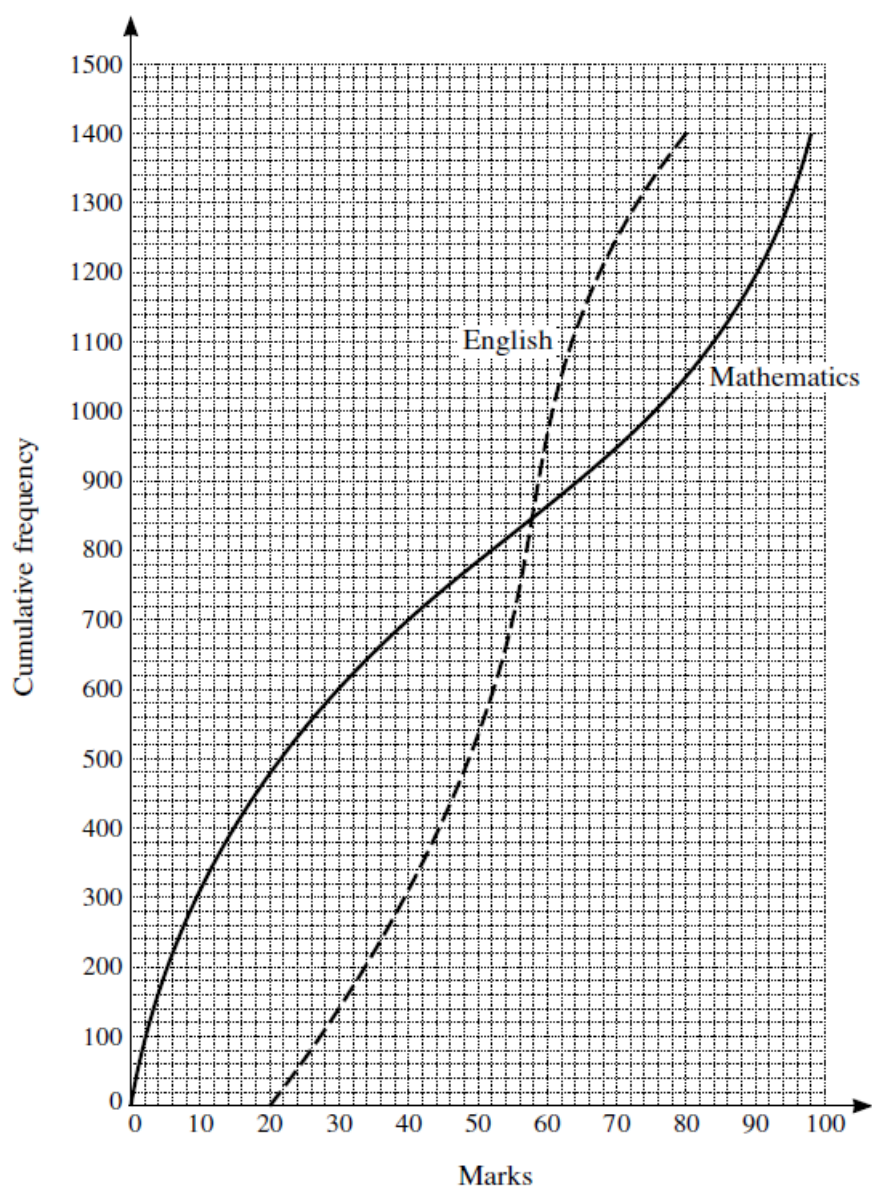
[2]

(iii) Calculate estimates of the mean and standard deviation of the daily rainfall in this village.

[6]

Q9 s2019_p61

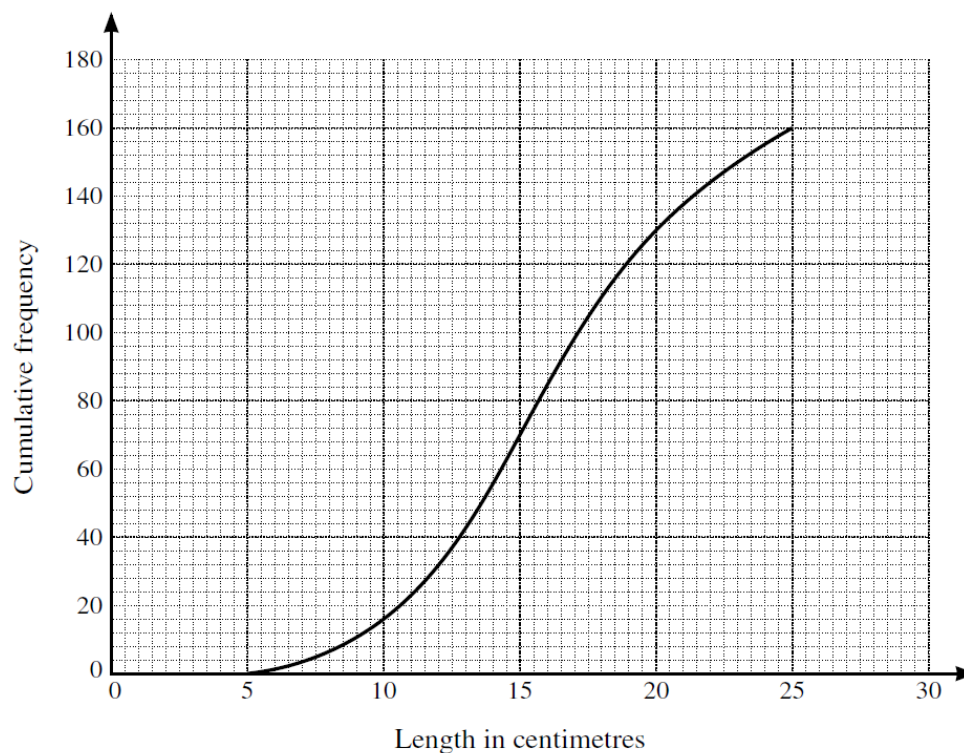
The Mathematics and English A-level marks of 1400 pupils all taking the same examinations are shown in the cumulative frequency graphs below. Both examinations are marked out of 100.



Use suitable data from these graphs to compare the central tendency and spread of the marks in Mathematics and English. [6]

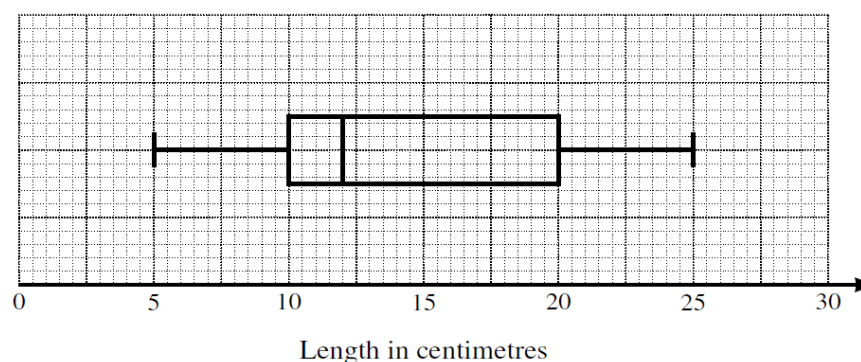
Q10 w2019_p61

Ransha measured the lengths, in centimetres, of 160 palm leaves. His results are illustrated in the cumulative frequency graph below.



- (i) Estimate how many leaves have a length between 14 and 24 centimetres. [1]
- (ii) 10% of the leaves have a length of L centimetres or more. Estimate the value of L . [2]
- (iii) Estimate the median and the interquartile range of the lengths. [3]

Sharim measured the lengths, in centimetres, of 160 palm leaves of a different type. He drew a box-and-whisker plot for the data, as shown on the grid below.



- (iv) Compare the central tendency and the spread of the two sets of data. [2]

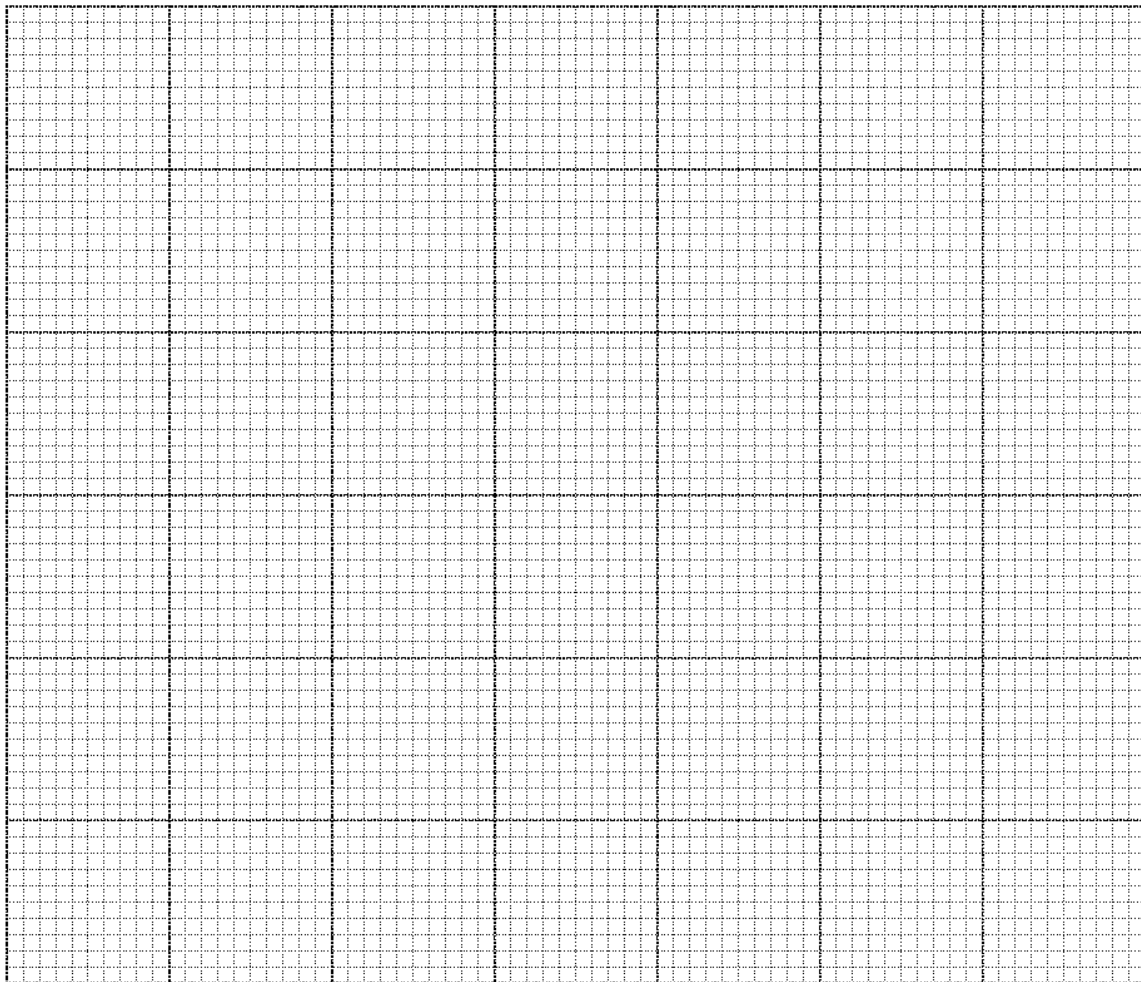
Q11 w2019_p63

Last Saturday, 200 drivers entering a car park were asked the time, in minutes, that it had taken them to travel from home to the car park. The results are summarised in the following cumulative frequency table.

Time (t minutes)	$t \leq 10$	$t \leq 20$	$t \leq 30$	$t \leq 50$	$t \leq 70$	$t \leq 90$
Cumulative frequency	16	50	106	146	176	200

(i) On the grid, draw a cumulative frequency graph to illustrate the data.

[2]



(ii) Use your graph to estimate the median of the data.

[1]

(iii) For 80 of the drivers, the time taken was at least T minutes. Use your graph to estimate the value of T .

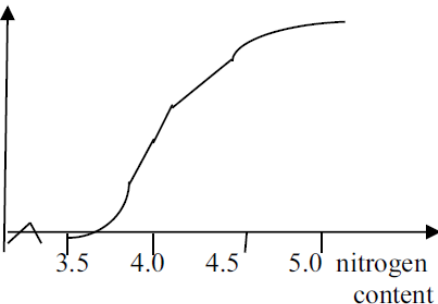
[2]

(iv) Calculate an estimate of the mean time taken by all 200 drivers to travel to the car park.

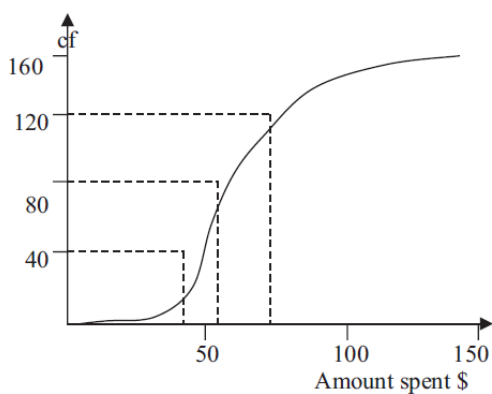
[4]

ANSWERS

Q1

(i)		B1 M1 A1 [3]	Uniform axes cf and nitrogen content labelled, at least 0 to 70 and 3.5 to 4.8 seen 5 points plotted correctly on graph paper <table border="1" data-bbox="979 598 1394 665"><tr><td>3.5</td><td>3.8</td><td>4.0</td><td>4.2</td><td>4.5</td><td>4.8</td></tr><tr><td>0</td><td>6</td><td>18</td><td>41</td><td>62</td><td>70</td></tr></table> All points correct and a reasonable curve (condone 1 missed point) or line segments.	3.5	3.8	4.0	4.2	4.5	4.8	0	6	18	41	62	70
3.5	3.8	4.0	4.2	4.5	4.8										
0	6	18	41	62	70										
(ii)	70 – their 55 = 15 = 21.4%	M1 A1 [2]	Subt a value > 41 from 70 (or $n/70$, $n < 29$) Correct ans, accept 18.5 – 22												
(iii)	median = 4.15	B1 [1]	Accept $4.1 < \text{median} < 4.2$, nfw												

Q2

(i)	cf 16, 56, 104, 130, 160	M1	Attempt at cf table (up to 160) no graph needed accept %cf but give final
		B1	linear scale minimum 0 to 160 and 0 to 120
		M1	Attempt to plot points at (30, 16), (50, 56), (70, 104), (90, 130), (140, 160) up to 2 errors can have a polygon
		A1 [4]	All points correct from their scale and joined up, with (0,0) as well
(ii)	median \$59	B1 ✓	accept 57–60 or fit their graph if used lb, midpts instead of ub or assume linear interpolation.
	IQR = 82 – 43 = \$39	M1	Subt a (sensible) LQ from a sensible UQ (generous)
		A1 ✓ [3]	Ans fit need a cf graph and UQ 80–84, LQ
(iii)	160 – 149 = 11 OR 115 is mid pt of last interval so # of shoppers is 30/2 = 15 (can be implied)	M1	41–46
		A1 [2]	Subtracting from 160 can be implied Correct answer accept 9–16
(iv)	mean = $(15 \times 16 + 40 \times 40 + 60 \times 48 + 80 \times 26 + 115 \times 30) / 160$ = $10250 / 160 = \$64.1 = \64.1	M1	Using $\Sigma xf / 160$ with mid-points
		A1 [2]	

Q3

(i)		B1	Horizontal axis from min of 140 to 190 and vertical axis from 0 to minimum of 60 and two CF graphs on the same set of axes.
		B1	Labels: CF; height (ht) in cm; girls; boys in correct places
		B1	CF graph going through (150, 0), (160, 20), (170, 43), (180, 55) and (190, 60)
		B1	CF graph going through (140, 0), (150, 12), (160, 33), (170, 50), (180, 60) [and (190, 60)]
		[4]	
(ii)	42 (± 1) shorter than 165. $(18(\pm 1))/60 \times 100$ $= 30\% (\pm 1.7\%)$	M1	Line or reading from 165 on their cf graph or subtracting from 60
		M1	
		A1	[3]
(iii)	can see which is taller; see which of boys or girls is more spread out	B1	any sensible comment in context
		[1]	

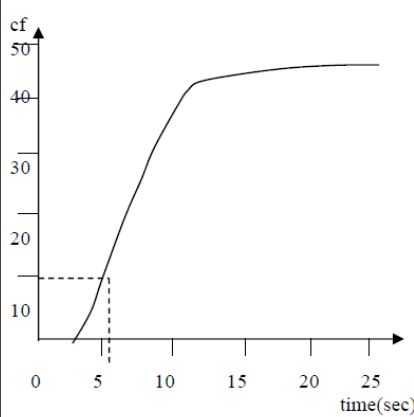
Q4

(i)	med = 3.2 $UQ = 3.65 \leq uq \leq 3.7$ $LQ = 2.55 \leq lq \leq 2.6$ $IQR = 1.05 \leq iqr \leq 1.15$	B1	Accept 3.2 ± 0.05
		M1	UQ – LQ, UQ greater than <i>their</i> ‘median’, LQ less than <i>their</i> ‘median’
		A1	Correct answer from both LQ and UQ in given ranges
	Total:	3	
(ii)	$134 - 24 = 110$	B1	Accept $108 \leq n \leq 112$, n an integer
	Total:	1	
(iii)	$200 - 12 = 188$ less than length l	M1	188 seen, can be implied by answer in range, mark on graph.
	$l = 4.5$ cm	A1	Correct answer accept $4.4 \leq l \leq 4.5$
	Total:	2	

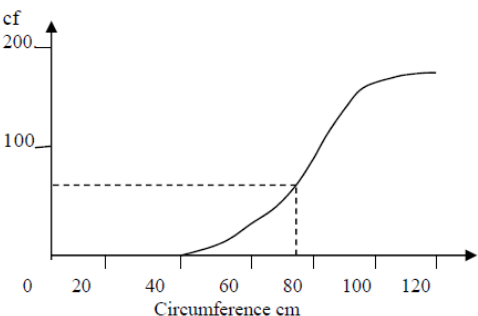
Q5

(i)	freq = fd $\times$ cw 10, 40, 120, 30	M1 A1	Attempt to multiply at least 3 fds by their 'class widths'										
Totals:		2											
(ii)	<table border="1"><tr><td>length</td><td>< 5</td><td>< 10</td><td>< 20</td><td>< 25</td></tr><tr><td>cf</td><td>10</td><td>50</td><td>170</td><td>200</td></tr></table>	length	< 5	< 10	< 20	< 25	cf	10	50	170	200	B1 B1 M1 A1	3 or more correct cfs heights on graph 10, 50, 170, 200 Labels correct cf and length(cm), linear scales from zero (allow 0.5 on horizontal axis) Attempt (at least three) at plotting at upper end points (either 5 or 5.5, 10 or 10.5 etc.) Starting at (0, 0) polygon or smooth curve increasing with plotted points at lengths 5, 10, 20 and 25
length	< 5	< 10	< 20	< 25									
cf	10	50	170	200									
Totals:		4											
(iii)	median = 14.2	B1	Median (accept 13.2 – 15.2)										
	'18.5' – '10'	M1	Subt their LQ from their UQ if reasonable from their graph										
	IQ range = 8.5	A1FT	Correct FT using LQ = 10 and UQ between 17.5 and 19.5										
	Totals:		3										
(iv)	mean = $(2.5 \times 10 + 7.5 \times 40 + 15 \times 120 + 22.5 \times 30) / 200$	M1	Using mid points (± 0.5) and their frequencies from 7(i) in correct formula										
	= 14	A1											
	Totals:		2										

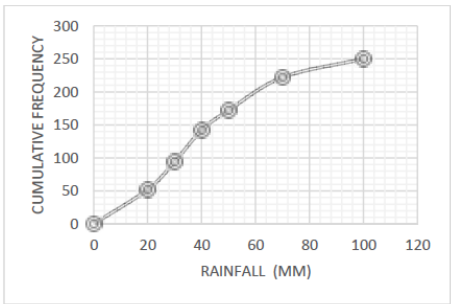
Q6

(i)	Points (5.5,10), (8.5,25), (11.5,42), (16.5,46), (25.5,48) 	B1	Correct cfs values seen listed, in or by table or on graph, 0 not required
		B1	Axes labelled "cumulative frequency" (or cf) and "time [or t etc.] (in) seconds (or sec etc.)". Linear scales – cf 0–48, time 2.5 – 25.5 (ignore <2.5 on time.) At least 3 values stated on each axis, but (0,0) can be implied without stating.
		B1	All points plotted accurately, (5, 10) etc. scores B0 . Curve or line segments drawn starting at (5.5,10) and passing within '1 scale unit' vertically and horizontally of plotted points
		3	
(ii)	$48 - 35 = 13$ $t = 6.5 \text{ sec}$	M1	Subt 35 (checked $\pm 1 \text{ mm}$ on graph) from 48 or 50,
		A1	$6 \leq \text{Ans} \leq 7$
		2	

Q7

(i)	points (50, 14), (80, 62), (100, 132), (120, 140) 	B1	Correct cfs values seen listed, in or by table or on graph, 0 not required
		B1	Axes labelled 'cumulative frequency' (or cf) and 'circumference [or cir or c etc.] (in) cm'. Linear scales – c.f. 0–140 circumference 40–120 (ignore <40 on circ.) At least 3 values stated on each axis, but (0,0) can be implied without stating.
		B1	All points plotted accurately
		3	
(ii)	$140 - 54 = 86$ Percentage = 61.4%	M1	Finding correct value from graph (checked $\pm 1 \text{ mm}$) or linear interpolation. Subtraction from 140 can be implied
		A1	$60.5\% \leq \text{Ans} \leq 64.5\%$
		2	

Q8

(i)		B1	Appropriate linear scales starting at (0,0), axes labelled cf and Rainfall, mm
		B1	Correct graph, points plotted at ucb, allow straight lines or curve
		2	
(ii)		M1	Read off from increasing graph at cf = 150
	42	A1	Correct answer ($41 \leq r \leq 43$)
		2	
(iii)	Frequencies 52, 42, 48, 30, 50, 28	B1	Correct frequencies
	Mean age = $(10 \times 52 + 25 \times 42 + 35 \times 48 + 45 \times 30 + 60 \times 50 + 85 \times 28) / 250$	B1	Correct midpoints (allow one error)
	= 9980/250	M1	Using $\Sigma fx/250$ with mid-points attempt, not cf, cw, lb, ub
	= 39.9(2) oe	A1	Correct answer
	Variance = $10^2 \times 52 + 25^2 \times 42 + 35^2 \times 48 + 45^2 \times 30 + 60^2 \times 50 + 85^2 \times 28 / 250 - \text{mean}^2$ = 539.59	M1	Attempt at variance using their midpoints and their mean
	$\sigma = 23.2$	A1	Correct answer for sd
		6	

Q9

Median Maths = 40	M1	Indication of finding medians, such as mark on graph or reference marks to 700 pupils, condone poor terminology such as 'mean'
Median English = 55	A1	Both values correct, condone $54 < \text{English} < 56$ but 54, 56 get A0
Median of English is larger than median of Maths	B1	Correct statement, median must be referenced within answer. No credit if statement references 'means'
Range Maths is 100 or IQ range Maths = $80 - 12 = 68$	M1	Evidence of finding either both ranges or both IQ ranges i.e. see a minus
Range English is 60 or IQ range English = $62 - 42 = 20$	A1	Both ranges or IQR correct
Maths marks have more spread than English marks	B1	Correct conclusion. Accept standard deviation but must see some figures
	6	

Q10

(i)	$156 - 55 = 99$	B1	$98 \leq \text{answer} < 100$
		1	
(ii)	90% of 160 = 144	M1	144 seen, may be marked on graph
	(L =) 22	A1	
		2	
(iii)	Median = 15.6 UQ = 18.8, LQ = 12.7	B1	$15.5 < \text{median} < 15.8$
	IQR = 18.8 – 12.7	M1	$18.5 < \text{UQ} < 19 - 12.5 < \text{LQ} < 13$
	6.1	A1	$6.0 \leq \text{IQR} \leq 6.2$
		3	
(iv)	The Median higher for Ransha (1st set of data)	B1	Any correct comparison of central tendency, must mention median
	IQR lower for Ransha (1st set of data)	B1	Any correct comparison of spread, must refer to IQR
		2	

Q11

(i)	Correct labels and scales	B1	Axes labelled ‘cumulative frequency’ (or cf) and ‘time (or t) [in] min(utes)’, linear scales from 0 to 90 and 0 to 200 with at least 3 values marked on each axis.
	7 correctly plotted points above upper boundaries joined in a curve or line segments	B1	(0, 0); (10, 16); (20, 50); (30, 106); (50, 146); (70, 176); (90, 200)
		2	
(ii)	29	B1	$28 \leq \text{median} \leq 30$
		1	
(iii)	120 seen	M1	For seeing 120 in a calculation or marked on the graph
	37	A1FT	$36 \leq \text{Ans} \leq 39$ or FT from <i>their</i> graph SC1 unsupported answer in range
		2	
(iv)	Frequencies 16 34 56 40 30 24	B1	Seen. Allow unsimplified
	Est. Mean = $\frac{5 \times 16 + 15 \times 34 + 25 \times 56 + 40 \times 40 + 60 \times 30 + 80 \times 24}{200}$	M1	At least 4 correct midpoints (5, 15, 25, 40, 60, 80) used in a calculation
	$\frac{7310}{200}$	M1	Summing products of <i>their</i> 6 mid-points (not lower or upper bound or class width) $\times$ <i>their</i> frequencies / 200 (or <i>their</i> $\sum f$), unsimplified
	36.55	A1	Accept 36.6
		4	